

**Amendments to the Specification:**

This listing of claims will replace all prior versions and listing of claims in the application:

**Listing of Claims:**

1. (original) A system for synchronization between a moving picture and a text-to-speech converter, comprising:

distributing means for receiving multi-media input information, transforming said multi-media input information into respective data structures, and distributing the respective data structures for further processing;

image output means for receiving image information of the distributed multi-media information and displaying the image information;

language processing means for receiving language texts of the distributed multi-media information, transforming the language texts into phoneme strings, and estimating and symbolizing prosodic information from the language texts

prosody processing means for receiving the prosodic information from said language processing means, and calculating values of prosodic control parameters;

synchronization adjusting means for receiving the prosodic control parameters from said prosody processing means, adjusting time durations for every phoneme for synchronization with the image information by using synchronization information of the distributed multi-media information, and inserting adjusted time durations into the prosodic control parameters;

signal processing means for receiving the processing results from said synchronization adjusting means and generating a synthesized speech; and

a synthesis unit database block for selecting required units for synthesis in accordance with a request from said signal processing means, and transmitting the required data to said signal processing means.

2. (currently amended) The system according to claim 1, wherein the multi-media information comprises:

the language texts, image information on moving picture, and synchronization information,

and wherein the synchronization information includes:

~~a text~~, information on a lip shape, information on image positions in the moving picture, and information on time durations.

3. (original) The system according to claim 2, wherein the information on the lip shape can be transformed into numerical values based on a degree of a down motion of a lower lip, up and down motion at a left edge of an upper lip, up and down motion at a right edge of the upper lip, up and down motion at a left edge of the lower lip, up and down motion at a right edge of the lower lip, up and down motion at a center portion of the upper lip, up and down motion at a center portion of the lower lip, a degree of protrusion of the upper lip, a degree of protrusion of the lower lip, a distance from the center of the lip to the right edge of the lip, and a distance from the center of the lip to the left edge of the lip,

and wherein the information on the lip shape is definable in a quantified and normalized pattern in accordance with the position and manner of articulation for each phoneme.

4. (original) The system according to claim 1, wherein said synchronization adjusting means comprises means for calculating time durations of phonemes within a text by using the synchronization information in accordance with a predicted lip shape determined by a position and manner of articulation for each phoneme within a text, a lip shape within the synchronization information, and time durations.

5. (new) The system of claim 2, wherein said synchronization information further includes text.